Complete Title
"Waiting to Die": Toxic Emissions and Disease Near the Denka Performance Elastomer Neoprene Facility in Louisiana's Cancer Alley

Authors
Ruhan Nagra[1]*, Robert Taylor[2], Mary Hampton[3], Lance Hilderbrand[4]**

*Corresponding author
**Current address: USC Equity Research Institute, 1149 South Hill Street, Suite H340, Los Angeles CA USA 90015

Running Title
Pollution and Disease Near Louisiana's Denka Plant

A preliminary version of this study was posted on the University Network for Human Rights website at: https://drive.google.com/file/d/1Ie93SHF-GrgFfN61PqwXrGh1Ay4lWqMD/view

_____

[1] University Network for Human Rights; Horgan House, Wesleyan University, 77 Pearl Street, Middletown CT USA 06457; ruhan@humanrightsnetwork.org; Telephone: 314.435.2377
[2] Concerned Citizens of St. John the Baptist Parish; 389 East 26th Street, Reserve LA USA 70084; btcnola@gmail.com; Telephone: 504.559.7304
[3] Concerned Citizens of St. John the Baptist Parish; 155 East 24th Street, Reserve LA USA 70084; hamptonmaryc@gmail.com; Telephone: 985.210.1976
[4] University Network for Human Rights; Horgan House, Wesleyan University, 77 Pearl Street, Middletown CT USA 06457; lancemh1@gmail.com; Telephone: 206.335.6929

1

Abstract

Background

Residents of census tract 708 in St. John Parish, Louisiana face the highest nationwide cancer risk from air pollution due to chloroprene emissions from the Denka Performance Elastomer facility. The University Network for Human Rights worked with residents of this predominantly Black community in Cancer Alley to design and implement a survey-based health study of the area. The study aimed to (1) assess the relationship between household proximity to the facility and reported illness, and (2) advance the advocacy objectives of the community.

Methods

The survey area consisted of households within a 2.5-kilometer radius of the Denka facility. 60% of the households within 1.5 kilometers of the facility ("Zone 1") and 20% of the households between 1.5 and 2.5 kilometers from the facility ("Zone 2") were randomly sampled. Survey implementers collected information on cancer diagnoses about all residents of each surveyed household. Information on chloroprene-linked medical symptoms was collected about respondents (those who took the survey) only.

Results

Cancer prevalence among the survey sample is: (1) significantly higher than what is considered likely using Monte Carlo simulations based on SEER prevalence data ($p = 0.0306$); and (2) associated with proximity to the facility, with significantly higher-than-likely prevalence in Zone 1 ($p = 0.0032$) and lower prevalence in Zone 2. Levels of medical symptoms among respondents are high and also associated with proximity to the facility.

Discussion

Our findings highlight the need for action to compel Denka to reduce chloroprene emissions to EPA-recommended limits.

Conclusion

Our findings are consistent with Cancer Alley communities' lived experiences of the debilitating health consequences of the area's industrial emissions. The burden of proof must shift to polluting industries.

2

Introduction

Cancer Alley and the Denka Neoprene Facility

Louisiana's heavily industrialized corridor between New Orleans and Baton Rouge has long been known as "Cancer Alley." Over 200 chemical plants and refineries are concentrated in this 130-mile stretch of land along the Mississippi River, mostly in or near historically Black communities where many residents can trace their lineage to ancestors who were enslaved in the area.[5] Since the late 1970s, many Cancer Alley residents have attributed cancer and other illness in their communities to toxic industrial pollution[6] and sought to use regulatory and legal challenges as well as grassroots struggle to compel industry to reduce emissions.[7]

In the last several years, Environmental Protection Agency (EPA) data have bolstered suspicions about the link between air pollution and negative health outcomes in Cancer Alley.[8] According to the most recent EPA National Air Toxics Assessment (NATA), seven of the ten U.S. census tracts with the highest cancer risk from air pollution are in Cancer Alley, including the tract with the highest nationwide risk—tract 708 in the town of Reserve in St. John the Baptist Parish.[9]

Nationally, the average estimated risk of developing cancer from air pollution is 32 per million people; in Louisiana's census tract 708, the estimated cancer risk from air pollution is 1,505 per million people—47 times the national average.[10] The vast majority of this risk, moreover, is attributed to a single chemical, chloroprene, emitted by the Denka Performance Elastomer neoprene facility. EPA attributes 85% (1,279 per million people) of the cancer risk from air pollution in census tract 708 to chloroprene emissions, 12% (187 per million people) to ethylene oxide emissions, and 3% (38 per million people) to all other pollutants.[11] The Denka facility is

[5] Trymaine Lee, "Cancer Alley: Big Industry, Big Problems," *MSNBC*, http://www.msnbc.com/interactives/geography-of-poverty/se.html.

[6] Barbara Allen, "Cradle of a Revolution? The Industrial Transformation of Louisiana's Lower Mississippi River," *Technology and Culture* 47 (2006): 115-16, doi:10.1353/tech.2006.0051.

[7] *Id* 116-17. In the Great Louisiana Toxics March of 1989, hundreds of Cancer Alley residents walked from Baton Rouge to New Orleans over a ten-day period. Thirty years later, in 2019, the Coalition Against Death Alley (CADA)—a coalition of community groups across Cancer Alley and their allies—marched from the town of Reserve to the state capitol in Baton Rouge, demanding environmental justice. Jamiles Lartey and Oliver Laughland, "'They've been killing us for too long': Louisiana residents march in coalition against 'death alley,'" *The Guardian*, May 30, 2019, https://www.theguardian.com/us-news/2019/may/30/toxic-america-louisiana-residents-march-against-polluting-plant.

[8] EPA's 2011 and 2014 National Air Toxics Assessment (NATA) data showed elevated cancer risks from air pollution in a number of Cancer Alley census tracts. According to the 2014 NATA, for example, of the 109 U.S. census tracts where the probability of developing cancer from air pollution is higher than EPA's upper limit of acceptable risk (100 per million people), 31 are in Cancer Alley. In addition, EPA's Risk-Screening Environmental Indicators (RSEI) model shows very high estimated levels of cancer-causing pollutants in Cancer Alley, according to a recent analysis. Lylla Younes, Al Shaw, and Claire Perlman, "In a Notoriously Polluted Area of the Country, Massive New Chemical Plants Are Still Moving In," *ProPublica*, October 30, 2019, https://projects.propublica.org/louisiana-toxic-air/.

[9] U.S. Environmental Protection Agency, "2014 National Air Toxics Assessment," August 2018, https://www.epa.gov/national-air-toxics-assessment/2014-nata-assessment-results#nationwide. We consider Cancer Alley to include the following eleven parishes (i.e. counties) of Louisiana: Ascension, East Baton Rouge, Iberville, Jefferson, Orleans, Plaquemines, St. Bernard, St. Charles, St. James, St. John the Baptist, and West Baton Rouge.

[10] *Id.*

[11] *Id.*

the only source of chloroprene emissions in St. John Parish[12] and the only producer of chloroprene and neoprene in the United States.[13]

The neoprene facility, owned by DuPont until its sale to Japanese company Denka Performance Elastomer in November 2015, has been pumping chloroprene into the neighboring Black community since 1969.[14] Residents of the community had long felt that there was too much illness in the area—far beyond what could be considered normal.[15] As one resident told us, "We're just sitting here, waiting to die."[16]

EPA's Integrated Risk Information System (IRIS) classified chloroprene as a "likely human carcinogen" in 2010. Reflecting this new IRIS assessment of chloroprene toxicity, the 2011 NATA (published in December 2015) estimated highly elevated cancer risk from air pollution near the Denka facility. Upon learning about EPA's estimate of their cancer risk in July 2016, residents of Reserve formed a community group called Concerned Citizens of St. John the Baptist Parish ("Concerned Citizens"). Concerned Citizens has demanded a significant reduction in chloroprene emissions from the Denka facility, such that air concentration of the chemical does not exceed 0.2 micrograms per cubic meter ($\mu g/m^3$)—the maximum chloroprene air concentration that would keep cancer risk from air pollution within EPA's "upper limit of acceptability" (100 per million people).[17] Concerned Citizens' ongoing struggle for environmental justice has gained increasing traction and national media coverage.[18]

In January 2017, Denka signed a voluntary agreement with the Louisiana Department of Environmental Quality to reduce its emissions.[19] Although chloroprene air concentrations have dropped since then, EPA's monitoring data have continued to show concentrations well in excess of 0.2 $\mu g/m^3$ in the neighborhoods around the Denka facility: in 2020, 35% of air samples

---

[12] Louisiana Department of Environmental Quality, "Annual Certified Emissions Data 1991-present," updated April 2020, https://www.deq.louisiana.gov/page/eric-public-reports.

[13] Jamiles Lartey and Oliver Laughland, "Cancer and chemicals in Reserve, Louisiana: the science explained," *The Guardian*, May 6, 2019, https://www.theguardian.com/us-news/2019/may/06/cancertown-chemicals-reserve-louisiana-science.

[14] Sharon Lerner, "The Plant Next Door," *The Intercept*, March 24, 2017, https://theintercept.com/2017/03/24/a-louisiana-town-plagued-by-pollution-shows-why-cuts-to-the-epa-will-be-measured-in-illnesses-and-deaths/.

[15] *Id.*

[16] "Gloria Dumas," YouTube video, 2:46, excerpts of interview conducted by University Network for Human Rights, posted by "University Network for Human Rights," August 8, 2019, https://www.youtube.com/watch?time_continue=63&v=F77MvXt6y88&feature=emb_logo.

[17] U.S. Environmental Protection Agency, "Preliminary Risk-Based Concentration Value for Chloroprene in Ambient Air," May 5, 2016, https://www.epa.gov/sites/production/files/2016-06/documents/memo-prelim-risk-based-concentrations050516.pdf.

[18] Sharon Lerner, "When Pollution Is a Matter of Life and Death," *New York Times*, June 22, 2019, https://www.nytimes.com/2019/06/22/opinion/sunday/epa-carniogens.html; Jamiles Lartey and Oliver Laughland, "'Almost every household has someone that has died from cancer,'" *The Guardian*, May 6, 2019, https://www.theguardian.com/us-news/ng-interactive/2019/may/06/cancertown-louisana-reserve-special-report; Rebecca Hersher, "After Decades Of Air Pollution, A Louisiana Town Rebels Against A Chemical Giant," *NPR*, March 6, 2018, https://www.npr.org/sections/health-shots/2018/03/06/583973428/after-decades-of-air-pollution-a-louisiana-town-rebels-against-a-chemical-giant; Victor Blackwell, Wayne Drash, and Christopher Lett, "Toxic tensions in the heart of 'Cancer Alley,'" *CNN*, October 20, 2017, https://www.cnn.com/2017/10/20/health/louisiana-toxic-town/index.html.

[19] Louisiana Department of Environmental Quality, "Administrative Order on Consent," January 6, 2017, https://www.deq.louisiana.gov/assets/docs/Denka/DENKA_AdministrativeOrderOnConsentAOCJan2017.pdf.

4

exceeded the 0.2 µg/m³ threshold and the mean chloroprene air concentration was 0.7 µg/m³—over three times the threshold (see Table 1).

Although EPA's estimates of air pollution-related cancer risk have been critical in elevating the longstanding concerns of Cancer Alley residents, these risk estimates have not compelled adequate action to protect human health. As discussed further below, although building upon risk estimates with health studies to determine observed levels of negative health outcomes is valuable, such studies should not be necessary to compel action to protect human health. Once EPA has determined that residents of certain areas may face unacceptably high health risks, strong and swift action is not only warranted but obligatory.[20]

Genesis and Goals of our Community-Engaged Research Project

The University Network for Human Rights (UNHR) is a nonprofit organization that works closely with communities affected by rights abuse to amplify and advance their struggles through community-led interdisciplinary research, documentation, and advocacy. The authors of this study—UNHR researchers and leaders of Concerned Citizens of St. John Parish—first met in fall 2017.[21] Concerned Citizens then convened several joint community meetings with UNHR researchers to discern residents' most pressing concerns and advocacy priorities. Residents discussed at length their anecdotal evidence of abnormally high levels of cancer and other illness in the community. Multiple people reported, for example, that in almost every household on the streets closest to the Denka facility, someone had cancer or had died of cancer. Residents felt that, to have an impact, this anecdotal evidence needed to be supplemented with quantitative data collected through a household health survey of the area near the plant.

After community members identified a survey-based household health study as one of their priorities, UNHR researchers began working closely with Concerned Citizens to develop a community-engaged research plan for implementation of the study. The goals of the study were: (1) to determine the overall health status of a large sample of residents living in the area of the Denka facility; (2) to assess the relationship between household proximity to the Denka facility and reported illness; and (3) to advance the advocacy objectives of Concerned Citizens by collecting and analyzing data that might be useful in the group's efforts to compel Denka to adhere to the EPA's 0.2 µg/m³ guideline for maximum chloroprene air concentration.

---

[20] According to the precautionary principle, one of the most significant developments in modern international environmental law, decisionmakers must take action to protect the environment and public health when there is scientific uncertainty. Principle 15 of the 1992 Rio Declaration on Environment and Development states, for example: "In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation." United Nations General Assembly, "Annex 1: Rio Declaration on Environment and Development," *Report of the United Nations Conference on Environment and Development*," August 12, 1992, https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151 _26_Vol.I_Declaration.pdf.

[21] At the time, Ruhan Nagra was a clinical instructor at Stanford Law School's Human Rights Clinic. She transitioned employment to the University Network for Human Rights in fall 2018 and has continued this work in that capacity.

The survey instrument focused on chloroprene-linked health outcomes, in particular, because (1) the vast majority of the cancer risk from air pollution near the Denka facility is due to chloroprene emissions; (2) these emissions can be attributed to the Denka facility because it is the only source of chloroprene emissions in St. John Parish; and (3) the study was motivated by community members' concern about their exposure to chloroprene, which EPA had recently brought to their attention following the release of the 2011 NATA.

Methods

Epidemiologists and statisticians at Stanford University provided input and guidance to ensure use of proper actuarial processes, study design methods, and survey implementation principles and techniques. As a field epidemiology investigation, the study was (1) initiated in response to what community members described as a public health crisis in the area near the Denka facility, and (2) conducted in the field, through survey-based collection of residents' health information.[22] Stanford University's Research Compliance Office has determined that no IRB review would have been required "[b]ecause the goal of this project was advocacy for a specific issue in a specific situation and not generalizable research."

Survey Instrument

To guide the development of our survey instrument (see Appendix 1), we used peer-reviewed studies based on similar household health surveys.[23] The survey instrument was designed to collect certain health and other information—including age, sex, part- or full-time residency status, cancer and other medical diagnoses, and child health—about all residents of a household. Additional information was collected about respondents (those who took the survey) only, including race/ethnicity and medical symptoms.

Many symptoms and diagnoses were included in the survey instrument because of their link to chloroprene exposure, according to EPA's *Toxicological Review of Chloroprene* (2010). Other symptoms and diagnoses were included after community members identified them as particular sources of concern in focus group sessions held in February 2018.

In addition to cancer diagnoses, the following chloroprene-linked health symptoms were included in the survey instrument: headache, dizziness, fatigue, shortness of breath, rapid heart rate, heart palpitations, chest pain, and irritation of the eyes, nose, throat, and skin.[24] In light of community members' particular concern about health impacts on children as well as evidence suggesting that children are more susceptible than adults to the toxic effects of chloroprene

---

[22] Richard A. Goodman, James W. Buehler, and Michael Gregg, "Field epidemiology defined," *Field Epidemiology* (2008): 3-15, doi:10.1093/acprof:oso/9780195313802.001.0001.
[23] Peter M. Rabinowitz, Ilya B. Slizovskiy, Vanessa Lamers, Sally J. Trufan, Theodore R. Holford, James D. Dziura, Peter N. Peduzzi, Michael J. Kane, John S. Reif, Theresa R. Weiss, and Meredith H. Stowe, "Proximity to Natural Gas Wells and Reported Health Status: Results of a Household Survey in Washington County, Pennsylvania," *Environmental Health Perspectives* 123 (2015): 21-26, https://doi.org/10.1289/ehp.1307732.
[24] U.S. Environmental Protection Agency, "Toxicological Review of Chloroprene," September 2010, https://www.epa.gov/sites/production/files/2016-10/documents/chloroprene.pdf. These conditions can affect people both short- and long-term following exposure to chloroprene.

exposure,[25] we also collected survey data on two specific symptoms in children: headaches and nosebleeds. Community members cited both of these symptoms as common in children who live and/or attend school in the area near the Denka facility. (Additionally, as noted, headaches are linked to chloroprene exposure.)

Finally, the survey instrument included questions on the frequency and strength of chemical odors in the area as well as residents' level of concern about pollution in their community.

A draft survey instrument was piloted with five residents of the area in February 2018 and modified accordingly for clarity and efficiency of data collection.

Study Design

The geographic scope of the study was the area within a 2.5-kilometer radius of the Denka facility. In Fig. 1, the outer circle circumscribes the entire survey area and the inner circle circumscribes the area within 1.5 kilometers of the facility. The facility—with a red dot at its center—can be seen at the center of the survey area. In the map on the right, grey dots represent households. Residents of the orange-colored census tract (708) face the nation's highest cancer risk from air pollution, according to EPA. Residents of the yellow-colored census tract (709) face the third-highest nationwide risk.

We ultimately surveyed 60% of households (267 out of 445) within the 1.5-kilometer radius of the plant ("Zone 1," as shown in Fig. 1) and 20% of households (271 out of 1,376) located between 1.5 and 2.5 kilometers from the plant ("Zone 2"). Households were randomly sampled. After obtaining addresses by census block online, we used a census batch geocoder to geocode the addresses. We determined that there are 445 total households in Zone 1 and 1,376 total households in Zone 2, according to 2010 census information. We designed our protocol to ensure that we would randomly survey at least 250 households in Zone 1 (56% of the Zone 1 total) and at least 250 households in Zone 2 (18% of the Zone 2 total). Assuming a survey response rate of approximately 50%, we used the R random number generator to generate a randomly-ordered list of all 445 households in Zone 1 (predicting that we would need to attempt to survey all 445 households to achieve our target number of 250 surveys in Zone 1). We also used the R random number generator to randomly select (and randomly order) 500 addresses in Zone 2 (predicting that we would need to attempt to survey at least 500 households to achieve our target number of 250 surveys in Zone 2). Once we had attempted to survey all 500 households on our Zone 2 list at least twice without reaching the target number of surveys (250), we generated a randomly-ordered list of all remaining households in Zone 2. To reach our target number of surveys for each zone, we attempted to survey almost every household in Zone 2 and every household in Zone 1. Thus, the survey response rate is equivalent to the percentage of households ultimately surveyed in each zone.

Study Protocol

One day prior to the start of survey implementation, a team of community members and UNHR researchers distributed flyers throughout the survey area. The flyers informed residents about the

---

[25] *Id.*

upcoming health survey, its goals, and the possibility that their household might be randomly selected for participation. The flyers also stated that residents' participation in the survey was entirely voluntary.

After undergoing intensive training and practice in survey implementation principles and techniques under the supervision of Stanford University experts, a team of fourteen Stanford undergraduates implemented the survey over nine days (March 22-30, 2018). The survey area was divided into seven geographic sub-areas for ease of survey implementation (that is, so that survey implementers could be assigned to a sub-area for a given period of time rather than having to walk long distances from household to household across the entire survey area). Survey implementers almost always worked in pairs. Each day, each pair of survey implementers was assigned to one of the seven geographic sub-areas and provided with a list of households in their sub-area. The list was randomized, but to reduce time spent walking between households, the route efficiency was optimized for each set of twenty addresses. Survey implementers attempted to survey each of the twenty route-optimized households twice before moving on to the next set of twenty. The following day, survey implementers made a third attempt to survey households that had been attempted twice the previous day, before moving on to the next set of households. Survey implementers generally did not visit a household more than three times. If a household member declined to participate in the survey, implementers did not attempt to survey that household again. Households were surveyed from approximately 9am to 7pm each day.

For each household surveyed, one household member (the "respondent") provided health and demographic information about themself and every other person living in the household. We use the term "residents" to refer to everyone for whom data were collected (that is, respondents plus all other household members).

Survey implementers obtained verbal informed consent from each respondent before proceeding. Upon encountering a potential respondent, survey implementers introduced themselves and conveyed the purpose of the survey. They explained that participation in the survey was voluntary; that, if the potential respondent chose to participate, neither their name nor the names of any of their household members would be recorded; that any information provided would remain strictly confidential and would not be shared outside our research team; and that the overall results of the study would be made public but no one's identity or identifying health information would be disclosed. If the respondent verbally consented to participate in the survey, one of the survey implementers asked the survey questions while the other recorded the respondent's answers on a paper survey.

Following completion of survey implementation, the data from each survey were manually entered into an electronic REDCap instrument.

Data Analysis

Monte Carlo Analyses of Cancer Prevalence

We used Monte Carlo simulations in RStudio to analyze our data on cancer prevalence among residents surveyed. We simulated a population in the United States with the same race, sex, and

8

age demographics as the survey sample. Using 10,000 simulations, we generated probability distributions of cancer prevalence in the simulated population based on the National Cancer Institute's 2015 Surveillance, Epidemiology, and End Results (SEER) data for 23-year cancer prevalence (see Appendix 2 for code abstract).[26] "Simulated" cancer prevalence refers to the probability distribution of outcomes generated by these 10,000 simulations. We then compared 23-year cancer prevalence in the survey sample ("observed" cancer prevalence) with the 23-year cancer prevalence values that are likely—based on SEER data broken down by race, sex, and age—in a demographically similar U.S. population (see Appendix 3 for the race/sex/age breakdown of the survey sample with corresponding SEER prevalence data for each demographic). We determined the probability (p-value) that a simulated population with the same race, sex, and age makeup as the survey sample would have a cancer prevalence as high or higher than that observed in the survey sample. We considered results significant when $P < 0.05$.[27]

For every resident in the survey sample, we had a corresponding resident—of the same race, sex, and age—in the simulated population. Each member of the simulated population was assigned a value of 0 (no cancer diagnosis in the previous 23 years) or 1 (one or more cancer diagnoses in the previous 23 years). The probability that a simulated resident in a certain race/sex/age group would be assigned 0 or 1 was based on SEER data. For example: According to SEER data, 23-year cancer prevalence among Black men between the ages of 60 and 69 is about 12.8%. In the simulated population, every Black male in his 60s was randomly assigned a value of 1 with probability $p = 12.8\%$ (otherwise, a value of 0 with probability $1-p = 87.2\%$). Each simulated resident was assigned a value of 0 or 1 in this manner, using the SEER cancer prevalence data for that resident's race/sex/age group. The process was then repeated 9,999 times to generate a total of 10,000 simulations. This enabled us to compare the observed cancer prevalence outcome in the survey sample to a distribution of cancer prevalence outcomes in the simulated population. Race, sex, and age were considered in our Monte Carlo analyses because SEER data are broken down by these three demographic variables. Other demographic variables (such as socioeconomic status) could not be considered because we lacked comparable national cancer prevalence data for other variables.

We ran Monte Carlo simulations for cancer prevalence in the overall survey area as well as by spatial zone. After separately determining cancer prevalence probabilities closer to the Denka facility (in Zone 1) and farther away from the facility (in Zone 2), we were able to determine whether or not there is an association between cancer prevalence among the survey sample and proximity to the Denka plant.

We ran Monte Carlo simulations both with and without a smoking exclusion criterion. This exclusion criterion removed all residents who live in households where anyone smokes on a daily basis. Since corresponding residents were also removed from the simulated population, the

---

[26] AM Noone, N Howlader, M Krapcho, D Miller, A Brest, M Yu, J Ruhl, Z Tatalovich, A Mariotto, DR Lewis, HS Chen, EJ Feuer, and KA Cronin (eds), "SEER Cancer Statistics Review, 1975-2015," *National Cancer Institute* (2018), https://seer.cancer.gov/archive/csr/1975_2015/results_merged/sect_02_all_sites.pdf.
[27] A lower p-value indicates a smaller probability that the observed difference is due to chance; in other words, the lower the p-value, the more likely that the observed difference is a true difference.

9

smoking exclusion criterion impacted the range of simulated outcomes as well as the survey outcome.

<u>Age-Adjusted Cancer Prevalence by Spatial Zone</u>

In addition to Monte Carlo analyses, crude survey data on cancer prevalence in each zone were age-adjusted to the U.S. Standard Population in the year 2000 so that the survey data by zone could be directly compared with SEER's national cancer prevalence (which is also age-adjusted to the 2000 U.S. Standard Population). Survey data were age-adjusted both with and without a smoking exclusion criterion.

<u>Health Symptoms and Pollution Data</u>

We did not use Monte Carlo simulations for health symptoms and pollution data because we lacked comparable national data by demographic group. Survey data on the following symptoms and pollution questions are presented by spatial zone: (1) headaches and nosebleeds in children; (2) chest pain and heart palpitations; (3) wheezing and difficulty breathing; (4) headaches, dizziness, and lightheadedness; (5) eye pain/irritation and watery eyes; (6) cough, sneezing, and sore/hoarse throat; (7) skin rash/irritation and itchy skin; (8) fatigue/lethargy; (9) chemical odors; and (10) concern about pollution.

Results

Analysis of EPA's Chloroprene Air Monitoring Data

Since 2016, EPA has collected chloroprene air concentration data from six monitoring sites surrounding the Denka facility.[28] Using this data, we calculated annual mean concentrations in two different ways (see Table 1): in our "lower bound" method, we replaced entries listed as "ND" (concentration not detected) with values of 0 $\mu g/m^3$ and kept all values below the method detection limit (0.0417 $\mu g/m^3$) as they are. In our "upper bound" method, we substituted 0.0417 $\mu g/m^3$ for each "ND" entry and for each value below 0.0417 $\mu g/m^3$.

In 2020, the maximum chloroprene air concentration detected was 22.6 $\mu g/m^3$, 113 times the 0.2 $\mu g/m^3$ threshold. The lower and upper bound mean concentrations that year—0.7175 $\mu g/m^3$ and 0.7349 $\mu g/m^3$, respectively—were both over three times the threshold. 35.4% of air samples collected in 2020 had a chloroprene concentration that exceeded 0.2 $\mu g/m^3$.

Analyses of Cancer Prevalence

Of the 1,640 total residents in the survey sample, eliminations from the data set were made as follows for the analyses of cancer prevalence: 98 part-time residents (defined as those who live in the household for only 1-5 days of the week, inclusive) were eliminated from the data set. 8 residents for whom we did not have all three pieces of necessary demographic information—race, sex, and age—were eliminated from the data set. 21 residents who reported a race/ethnicity

---

[28] U.S. Environmental Protection Agency, "DENKA Air Monitoring Summary Sheet," updated September 26, 2020, https://www.epa.gov/sites/production/files/2020-10/documents/r6_summary_through_september_26_2020.pdf.

for which there is no SEER analogue (and therefore no comparable national cancer prevalence statistic) were eliminated from the data set. Finally, since we used SEER's 23-year cancer prevalence statistics, we eliminated the 6 residents whose only cancer diagnosis happened in 1994 or earlier (more than 23 years prior to the health survey).

After all eliminations, the numbers of residents included in the cancer prevalence analyses were **777 in Zone 1** (from 262 households) and **730 in Zone 2** (from 263 households), for a total of **1,507** (from 525 households).

Although race information was collected for respondents only, we assumed—for purposes of the cancer prevalence analyses only—that all residents of a household shared the race of the respondent. If a particular respondent was eliminated from the data set (due to one of the above elimination criteria), all members of the respondent's household were eliminated from the data set as well (since the other household members' race depended on the respondent's race).

<u>Monte Carlo Analyses of Cancer Prevalence Across Survey Area</u>

In a probability distribution of 10,000 simulations, the median value for 23-year cancer prevalence in a population with the same race, sex, and age demographics as the survey sample was 4.4% (Fig. 2). In other words, half of the simulations yielded cancer prevalence values lower than 4.4% and half of the simulations yielded cancer prevalence values higher than 4.4%. The median is therefore an approximation of the cancer prevalence outcome that is most likely in a simulated population with the same demographic makeup as the survey sample.[29] In Fig. 2, the median is represented by the dotted vertical line in the distribution.

The percentage of survey residents who reported at least one cancer diagnosis in the previous 23 years ("observed cancer prevalence") was 5.4%, significantly higher than indicated by Monte Carlo simulations based on SEER prevalence data (p = 0.0343) (Fig. 2). This p-value indicates the probability that a simulated population with the same demographic makeup as the survey sample would have a cancer prevalence greater than or equal to that of the survey sample. In Fig. 2, the survey sample cancer prevalence is represented by the solid red vertical line in the distribution. The greater the distance between the solid red line (survey sample cancer prevalence) and the dotted line (approximation of most likely cancer prevalence), the more unusual the cancer prevalence in the survey sample.

When the smoking exclusion criterion was applied, the median value for cancer prevalence in the probability distribution for the simulated population was 4.5% (Appendix 4). The percentage of survey residents who reported a cancer diagnosis in the previous 23 years was 5.4%, significantly higher than indicated by Monte Carlo simulations based on SEER prevalence data (p = 0.0306) (Appendix 4).

---

[29] The table in Fig. 2 also provides: (1) Minimum, i.e., the lowest cancer prevalence value in the probability distribution; (2) 1st Quartile, i.e., the cancer prevalence value at which 25% of the simulations yielded lower values and 75% of the simulations yielded higher values; (3) 3rd Quartile, i.e., the cancer prevalence value at which 75% of the simulations yielded lower values and 25% of the simulations yielded higher values; (4) Maximum, i.e., the highest cancer prevalence value in the probability distribution.

11

Monte Carlo Analyses of Cancer Prevalence by Spatial Zone

In probability distributions of 10,000 simulations by spatial zone, the median value for cancer prevalence in Zone 1 was 4.6% and the median value for cancer prevalence in Zone 2 was 4.4% (Fig. 3). In other words, in Zone 1, half of the simulations yielded cancer prevalence values lower than 4.6% and half of the simulations yielded cancer prevalence values higher than 4.6%; and in Zone 2, half of the simulations yielded cancer prevalence values lower than 4.4% and half of the simulations yielded cancer prevalence values higher than 4.4%. The median is therefore an approximation of the cancer prevalence outcome that is most likely in a simulated population with the same demographic makeup as the survey sample for each zone.[30] In Fig. 3, the red distribution shows the range of cancer prevalence values likely for a simulated population with the same demographic makeup as the Zone 1 survey sample, and the blue distribution shows the range of cancer prevalence values likely for a simulated population with the same demographic makeup as the Zone 2 survey sample. Because there is not a significant difference in the range of simulated cancer prevalence outcomes for Zone 1 and Zone 2, the two distributions overlap significantly. The median for Zone 1 is represented by the dotted red vertical line, and the median for Zone 2 is represented by the dotted blue vertical line.

The percentage of survey residents in Zone 1 who reported a cancer diagnosis was 6.7%, significantly higher than indicated by Monte Carlo simulations based on SEER prevalence data (p = 0.0033) (Fig. 3). This p-value indicates the probability that a simulated population with the same demographic makeup as the Zone 1 survey sample would have a cancer prevalence greater than or equal to that of the survey sample. The percentage of survey residents in Zone 2 who reported a cancer diagnosis was 4.1% (Fig. 3). In Fig. 3, Zone 1 cancer prevalence is represented by the solid red vertical line, and Zone 2 cancer prevalence is represented by the solid blue vertical line. The greater the distance between the solid line (survey sample cancer prevalence for zone) and dotted line of corresponding color (approximation of most likely cancer prevalence for zone), the more unusual the survey sample cancer prevalence for that zone.

When the smoking exclusion criterion was applied, the median value for cancer prevalence in the Zone 1 probability distribution was 4.6% and the percentage of Zone 1 survey residents who reported a cancer diagnosis was 7.0%, significantly higher than indicated by Monte Carlo simulations based on SEER prevalence data (p = 0.0032) (Appendix 4). The median value in the Zone 2 probability distribution was 4.5% and the percentage of Zone 2 survey residents who reported a cancer diagnosis was 4.3% (Appendix 4).

Age-Adjusted Cancer Prevalence by Spatial Zone

Age-adjusted cancer prevalence among residents surveyed in Zone 1 was 5.0139%, 44% higher than SEER's age-adjusted national cancer prevalence of 3.4851%. When the smoking exclusion criterion was applied, age-adjusted Zone 1 prevalence was 5.1421%, 48% higher than the

---

[30] The table in Fig. 3 also provides: (1) Minimum, i.e., the lowest cancer prevalence value in each probability distribution; (2) 1st Quartile, i.e., the cancer prevalence value for each distribution at which 25% of the simulations yielded lower values and 75% of the simulations yielded higher values; (3) 3rd Quartile, i.e., the cancer prevalence value for each distribution at which 75% of the simulations yielded lower values and 25% of the simulations yielded higher values; (4) Maximum, i.e., the highest cancer prevalence value in each probability distribution.

national prevalence of 3.4851%. Age-adjusted cancer prevalence among residents surveyed in Zone 2 was 3.5308%. When the smoking exclusion criterion was applied, age-adjusted Zone 2 prevalence was 3.5112%.

Race/Ethnicity, Health Symptoms, and Pollution Data

The race/ethnicity, health symptoms, and pollution data presented below were collected for survey respondents only, with the exception of data pertaining to children in the household. After part-time respondents were eliminated from the data set, the sample size for race/ethnicity, symptoms, and pollution data was **263 in Zone 1** and **259 in Zone 2** (a total of **522**). Data on headaches and nosebleeds in children were collected from survey respondents, who were asked about the health of any children in their households. After part-time children were eliminated from the data set, the sample size for child health data was **186 in Zone 1** and **220 in Zone 2** (a total of **406**).

Race/Ethnicity Data

The overwhelming majority of respondents in the survey area (80.7%) identified as Black. 15.7% of respondents identified as white, and 3.6% identified as another race/ethnicity or did not provide race/ethnicity information. Black respondents were not distributed evenly throughout the survey area. In Zone 1, a higher proportion of respondents identified as Black than in Zone 2 (93.2% versus 68.0%). Conversely, 4.9% of Zone 1 respondents and 26.6% of Zone 2 respondents identified as white. 1.9% of Zone 1 respondents and 5.4% of Zone 2 respondents identified as another race/ethnicity or did not provide race/ethnicity information.

Health Symptoms Data

Over 40% of children in households surveyed in Zone 1 (40.3%) reportedly suffer from headaches. This proportion dropped to 28.6% in Zone 2. Over one-fifth of children in households surveyed in Zone 1 (21%) reportedly suffer from nosebleeds. This proportion dropped slightly in Zone 2, to 18.2%. Nearly 40% of Zone 1 respondents (37.3%) reported that they experienced chest pain, heart palpitations, or both at least one day per week in the past month. This proportion dropped to 27.8% in Zone 2. Approximately one-third of Zone 1 respondents (33.5%) reported that they experienced wheezing and/or difficulty breathing at least two days per week in the past month. This proportion dropped to 24.3% in Zone 2. Over half of Zone 1 respondents (50.6%) reported that they experienced headaches, dizziness, and/or lightheadedness at least two days per week in the past month. This proportion dropped to 37.5% in Zone 2. Nearly half of Zone 1 respondents (44.5%) reported that they experienced eye pain/irritation and/or watery eyes at least two days per week in the past month. This proportion was roughly the same in Zone 2 (43.6%). Over 40% of Zone 1 respondents (41.1%) reported that they experienced cough, sneezing, and/or sore/hoarse throat at least four days per week in the past month. This proportion dropped to 33.6% in Zone 2. Over one-third of Zone 1 respondents (34.6%) reported that they experienced skin rash/irritation and/or itchy skin at least two days per week in the past month. This proportion dropped slightly in Zone 2, to 30.5%. Nearly 30% of Zone 1 respondents (29.3%) reported that they experienced fatigue/lethargy at least four days per week in the past month. This proportion dropped to 22.8% in Zone 2.

13

<u>Pollution Data</u>

Approximately half of Zone 1 respondents (49.4%) reported that they smell chemical odors while inside their homes "at least a few times per month." This proportion dropped to 31.3% in Zone 2. Over half of Zone 1 respondents (51.7%) reported that they smell chemical odors while outside their homes "at least a few times per week." This proportion dropped to 42.1% in Zone 2. Over three-fourths of Zone 1 respondents (76.4%) reported that they smell chemical odors while outside their homes "at least a few times per month." This proportion dropped to 67.2% in Zone 2. 84.0% of Zone 1 respondents reported that they are "extremely concerned" about pollution in their community. This proportion dropped to 63.7% in Zone 2.

Discussion

To our knowledge, this is the first study conducted in Cancer Alley that evaluates the potential link between household proximity to a particular industrial facility and reported adverse health outcomes. Our analysis yielded three major findings. First, cancer prevalence among the survey sample is significantly higher than what is considered likely using Monte Carlo simulations based on SEER prevalence data. Second, cancer prevalence among the survey sample is associated with proximity to the Denka facility, with significantly higher-than-likely prevalence in the zone closer to the facility and lower prevalence in the zone further from the facility. Third, levels of chloroprene-linked health symptoms among the survey sample—including among children—are high and also associated with proximity to the Denka facility.

Across the survey area as a whole, cancer prevalence among residents surveyed is significantly higher than what is considered likely for a U.S. population with the same race, sex, and age makeup. Removing residents who live in households where anyone smokes on a daily basis does not alter this result. When cancer prevalence among the survey sample is analyzed by spatial zone, prevalence in the zone closer to the Denka facility (Zone 1) is more statistically significant (with a p-value ten times lower) than prevalence in the survey area as a whole. Prevalence in Zone 1 is higher than prevalence in Zone 2, further from the facility. Again, applying the smoking exclusion criterion does not alter this result.

Our findings on other adverse health outcomes linked to chloroprene exposure show that high proportions of respondents regularly experience cardiac symptoms, difficulty breathing, headaches, eye irritation, respiratory symptoms, skin irritation, and fatigue. In virtually every case, respondents who live closer to the Denka facility (Zone 1) are affected in higher proportions than respondents who live further away (Zone 2).

Our findings on child health show that over 40% of children in surveyed households in Zone 1 suffer from headaches, an outcome linked to short- and long-term chloroprene exposure. Since the beginning of their struggle for environmental justice, Concerned Citizens of St. John Parish has advocated for the health and well-being of the children in their community. In particular,

14

Fifth Ward Elementary School—located less than a third of a mile from the Denka facility—has been a focal point of activism.[31]

A strength of the study was the random sampling design, which reduced the possibility of selection bias. Race data from survey samples in Zones 1 and 2 were representative of the respective larger areas: according to American Community Survey data, Zone 1 is 95% Black and 5% white (compared to 93% Black and 5% white in the survey sample) and Zone 2 is 71% Black and 27% white (compared to 68% Black and 27% white in the survey sample).[32] Additional strengths of the study included the spatial analysis of the data, i.e., the use of geographic zones by proximity to the facility; the consideration of confounding variables such as smoking, age, sex, and race; the value of field epidemiology, i.e., data collection in the field to investigate concerns about community health; and the strong partnership and relationship of trust between researchers and community members, which facilitated the design of a robust survey instrument (including through the use of focus groups) and collection of a large amount of data. Survey respondents were neither aware that the study design relied on the use of geographic zones nor aware of the zone in which their residence was located, reducing the possibility of awareness bias.

A limitation of the study was the reliance on self-reported health information provided by a single household member about all members of the household. On the one hand, respondents may have underreported other household members' health conditions. On the other hand, awareness bias in respondents who were concerned about air pollution, their own health, or household members' health may have increased reporting of adverse health outcomes. Other limitations included the use of only two comparison groups, limiting the ability to conduct statistical tests; the lack of reliable statistics to enable robust comparison of symptoms data; and potential confounding factors that were not considered, such as inclusion of multiple household members who share an indoor environment and may share genetics. In addition, our use of proximity to the facility was an indirect measure of exposure to air emissions; more precise measures of exposure include air monitoring and biomonitoring of individuals. Finally, stigma associated with illness—especially cancer—in the community may have led to a nonresponse bias that favored healthier individuals and households.

None of our findings came as a surprise to community members; rather, the study findings were consistent with community members' lived experiences. Community members view the health study as a useful tool to advance their struggle for clean air. Simultaneously—five years after discovering that they face the highest likelihood in the country of developing cancer from air pollution—residents are weary of hearing and reading about adverse health outcomes and pollution in their community and believe that it is long past time for action. More than sufficient evidence of chloroprene toxicity and community suffering has been collected to justify action; now, the state must compel Denka Performance Elastomer to reduce emissions so that chloroprene air concentration does not exceed EPA's maximum guideline of 0.2 $\mu g/m^3$.

---

[31] Nick Reimann, "St. John School Board panel suggests study on moving students from school near chemical plant," *The Advocate*, August 27, 2019, https://www.nola.com/news/education/article_275fc7d2-c83a-11e9-8fa9-87f1f4a3225a.html.

[32] The EPA's EJSCREEN tool was used to generate maps of Zones 1 and 2 and download 2013-2017 American Community Survey data for each zone.

Conclusion

EPA's estimate of cancer risk alone should have been enough to warrant swift and decisive action. As valuable as they are, health studies such as this one should not be necessary to compel decisionmakers to act to protect public health. Consistent with the precautionary principle in environmental science—which maintains that "when an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically"—action to protect public health in St. John should be taken on the basis of EPA's estimate of cancer risk in the parish.[33] Producing definitive scientific proof of a cause-and-effect relationship between chloroprene emissions and cancer in the area of the Denka facility would be virtually impossible—a feature of scientific uncertainty that polluting industries have long exploited to maintain their potentially toxic activities. Communities across Cancer Alley should not have to bear the burden of proof to achieve environmental justice. It is long past time for this burden to shift to Denka and other industries that are threatening human and environmental health.

---

[33] David Kriebel, Joel Tickner, Paul Epstein, John Lemons, Richard Levins, Edward L. Loechler, Margaret Quinn, Ruthann Rudel, Ted Schettler, and Michael Stoto, "The Precautionary Principle in Environmental Science," *Environmental Health Perspectives* 109 (2001): 871, https://doi.org/10.1289/ehp.01109871.

Acknowledgements

Author Disclosure Statement

The authors have no conflicts of interests or financial ties to disclose.

References

Allen, Barbara L. "Cradle of a Revolution? The Industrial Transformation of Louisiana's Lower Mississippi River." *Technology and Culture* 47 (2006): 112–19. doi:10.1353/tech.2006.0051.

Kriebel, David, Joel Tickner, Paul Epstein, John Lemons, Richard Levins, Edward L. Loechler, Margaret Quinn, Ruthann Rudel, Ted Schettler, and Michael Stoto. "The Precautionary Principle in Environmental Science." *Environmental Health Perspectives* 109 (2001): 871–76. https://doi.org/10.1289/ehp.01109871.

Lartey, Jamiles, and Oliver Laughland. "Cancer and chemicals in Reserve, Louisiana: the science explained." *The Guardian*, May 6, 2019. https://www.theguardian.com/us-news/2019/may/06/cancertown-chemicals-reserve-louisiana-science.

Lartey, Jamiles, and Oliver Laughland. "'They've been killing us for too long': Louisiana residents march in coalition against 'death alley.'" *The Guardian*, May 30, 2019. https://www.theguardian.com/us-news/2019/may/30/toxic-america-louisiana-residents-march-against-polluting-plant.

Lee, Trymaine. "Cancer Alley: Big Industry, Big Problems." *MSNBC*. http://www.msnbc.com/interactives/geography-of-poverty/se.html.

Lerner, Sharon. "The Plant Next Door." *The Intercept*, March 24, 2017. https://theintercept.com/2017/03/24/a-louisiana-town-plagued-by-pollution-shows-why-cuts-to-the-epa-will-be-measured-in-illnesses-and-deaths/.

Noone AM, N Howlader, M Krapcho, D Miller, A Brest, M Yu, J Ruhl, Z Tatalovich, A Mariotto, DR Lewis, HS Chen, EJ Feuer, and KA Cronin (eds). "SEER Cancer Statistics Review, 1975-2015." *National Cancer Institute* (2018). https://seer.cancer.gov/archive/csr/1975_2015/results_merged/sect_02_all_sites.pdf.

Rabinowitz, Peter M., Ilya B. Slizovskiy, Vanessa Lamers, Sally J. Trufan, Theodore R. Holford, James D. Dziura, Peter N. Peduzzi, Michael J. Kane, John S. Reif, Theresa R. Weiss, and Meredith H. Stowe. "Proximity to Natural Gas Wells and Reported Health Status: Results of a Household Survey in Washington County, Pennsylvania." *Environmental Health Perspectives* 123 (2015): 21-26. https://doi.org/10.1289/ehp.1307732.

Reimann, Nick. "St. John School Board panel suggests study on moving students from school near chemical plant." *The Advocate*, August 27, 2019. https://www.nola.com/news/education/article_275fc7d2-c83a-11e9-8fa9-87f1f4a3225a.html.

U.S. Environmental Protection Agency. "2014 National Air Toxics Assessment," August 2018. https://www.epa.gov/national-air-toxics-assessment/2014-nata-assessment-results#nationwide.

U.S. Environmental Protection Agency. "DENKA Air Monitoring Summary Sheet," updated September 26, 2020. https://www.epa.gov/sites/production/files/2020-10/documents/r6_summary_through_september_26_2020.pdf.

U.S. Environmental Protection Agency. "Preliminary Risk-Based Concentration Value for Chloroprene in Ambient Air," May 5, 2016. https://www.epa.gov/sites/production/files/2016-06/documents/memo-prelim-risk-based-concentrations050516.pdf.

U.S. Environmental Protection Agency. "Toxicological Review of Chloroprene," September 2010. https://www.epa.gov/sites/production/files/2016-10/documents/chloroprene.pdf.

Table 1. Summary Statistics of EPA's Chloroprene Air Monitoring Data

Year	Maximum Concentration Detected (μg/m³)	Mean Concentration (Lower Bound) (μg/m³)	Mean Concentration (Upper Bound) (μg/m³)	Proportion of Samples Greater Than 0.2 μg/m³
2016	153.0	7.3289	7.3387	68.6%
2017	151.0	3.7076	3.7190	53.5%
2018	98.7	2.1262	2.1393	47.8%
2019	27.2	1.1558	1.1737	46.5%
2020	22.6	0.7175	0.7349	35.4%

Fig. 1. Maps of survey area

Census Tracts (**bold** are included within survey area):

070800   **071100**   **070500**   070300

070900   **070700**   071000   070400

lon: -90.524
lat: 30.056

Fig. 2. Simulated and observed 23-year cancer prevalence
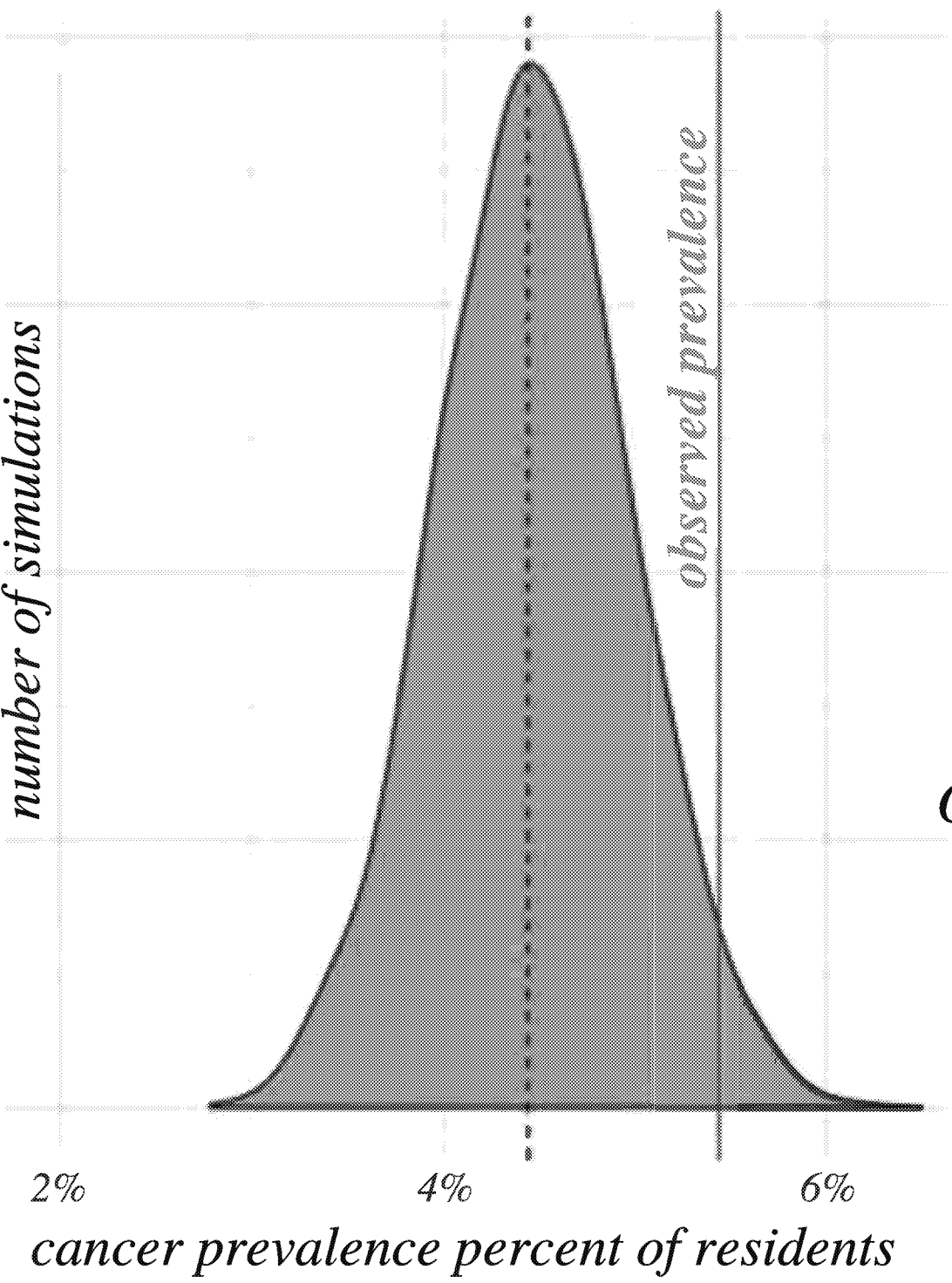
Simulated Cancer Prevalence

Statistic	Value
Minimum	2.8%
1st Quartile	4.1%
Median	4.4%
3rd Quartile	4.8%
Maximum	6.5%

Observed cancer prevalence: **5.4%**

p-value: **0.0343**

Fig. 3. Simulated and observed 23-year cancer prevalence by zone
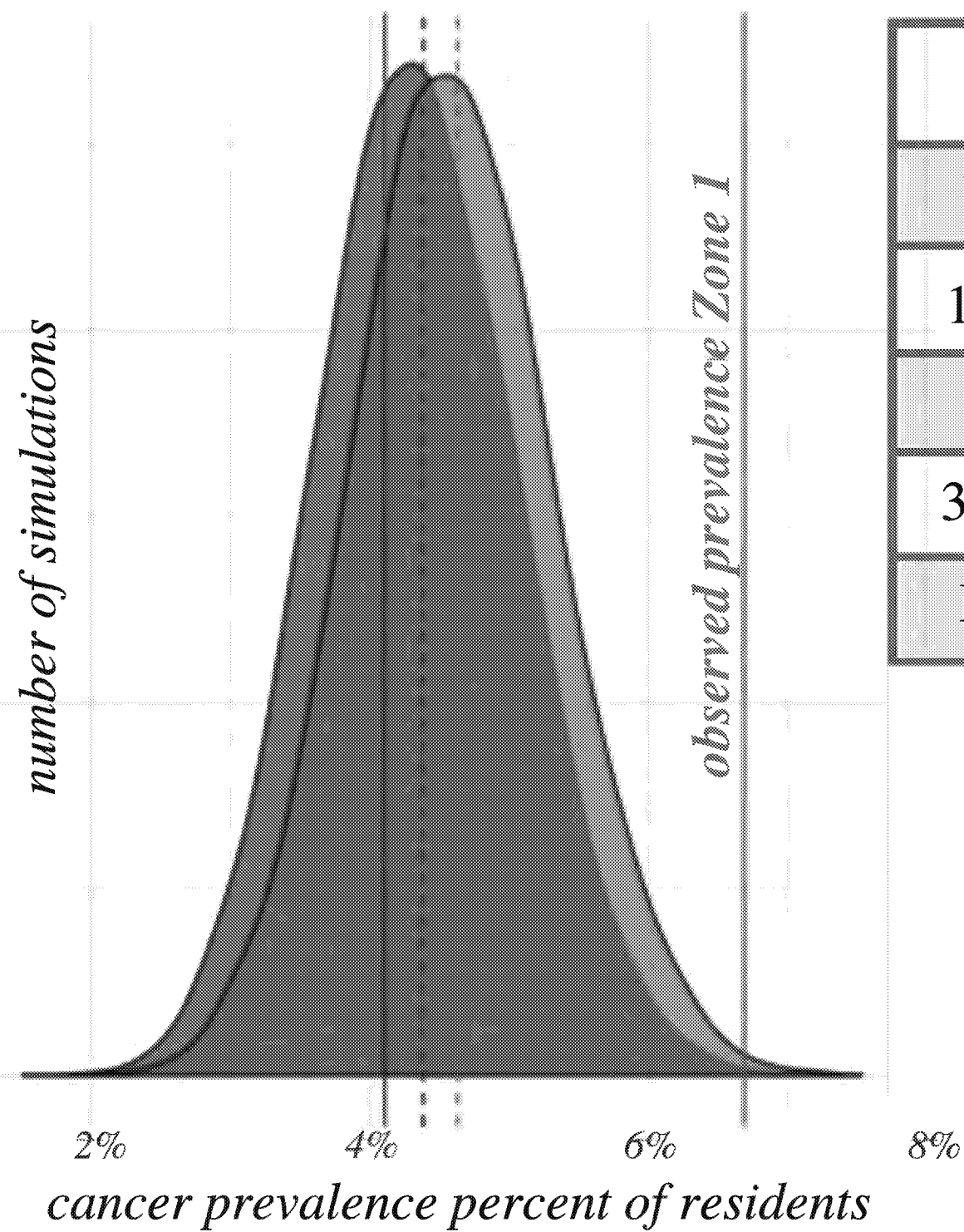
Simulated Cancer Prevalence

Statistic	Zone 1 value	Zone 2 value
Minimum	2.2%	1.5%
1st Quartile	4.1%	3.8%
Median	4.6%	4.4%
3rd Quartile	5.1%	4.8%
Maximum	7.3%	7.5%

Observed cancer prevalence:
Zone 1: **6.7%**; Zone 2: **4.1%**
p-value:
Zone 1: **0.0033**; Zone 2: **0.6527**

Stanford University

Participant ID#: _____

Community Health Survey
St. John the Baptist Parish

Data Collector 1: _____
Data Collector 2: _____
Date: _____ Time: _____

First, I'd like to ask some basic questions about you and each member of your household. We won't record names, just first initials.

Initial	Age (years)	Sex (M/F)	Blood relative? (Y/N)	Part- or full-time resident*	School (if 18 or under)
			N/A (self)		

*A part-time resident is someone who lives in the household for 1-5 days of the week (inclusive)

Now I'm going to ask you some basic questions about yourself, where you live, and where you used to live.

1. What is your race/ethnicity? (check all that apply)	☐ Asian ☐ Black or African American ☐ Hispanic/Latina/Latino/Latinx ☐ Native American ☐ Native Hawaiian and Other Pacific Islander ☐ White ☐ Other: _____
2. How long have you lived in this home?	☐ Less than one year ☐ ____ year(s)
3. Where did you live before moving to this home? (*city and state*)	
4. How long did you live in your previous home?	☐ Less than one year ☐ ____ year(s)

5. Now I'm going to ask you some questions about your medical history and the medical histories of others in your household. I'm going to go through a list of medical conditions. For each medical condition, I'll ask you whether a doctor or another health care provider has ever told you or anyone else in your household that you or they have the condition, and if so, what year you or they were told that. (*For the survey respondent, write yes or no, and year if relevant and known. For household members, provide the initial of every household member who has received the diagnosis, as well as year of diagnosis, if known.*)

	Yourself? (*yes/no, year*)	Household members? (*if yes, initial and year*)
a. ADHD?		
b. Allergies?		
c. Asthma?		
d. Anemia?		
e. Birth defects? Which one(s):		
f. Bronchitis?		
g. Congestive heart failure?		
h. Diabetes, other than during pregnancy?		
i. Heart disease?		
j. High blood pressure?		
k. Hyperthyroidism?		
l. Hypothyroidism?		
m. Learning difficulties?		
n. Nodules or a mass on the liver?		
o. Nodules or a mass on the lung(s)?		
p. Rapid pulse or rapid heartrate?		
q. Sinus infection?		

6. Now I'm going to ask about all members of your household and whether or not they had cancer, beginning with yourself. Please tell me the <u>month and year</u> of diagnosis, if possible. If members of your household had cancer and died, we will ask you about them afterward.

Type of cancer	Yourself? (yes/no, month & year)	Household members? (if yes, initial and month & year)
a. Bladder cancer		
b. Brain cancer		
c. Breast cancer		
d. Colon cancer		
e. Esophageal cancer		
f. Kidney cancer		
g. Leukemia		
h. Liver cancer		
i. Lung cancer		
j. Lymphoma		
k. Melanoma		
l. Oral cancer		
m. Ovarian cancer		
n. Pancreatic cancer		
o. Prostate cancer		
p. Sarcoma		
q. Skin cancer		
r. Spleen cancer		
s. Thyroid cancer		
t. Uterine cancer		
u. Other (specify):		

Now I'm going to ask you a few questions about others in your household and family.

7a. Has anyone in this household had cancer and died in the past 20 years?	☐ Yes If YES, who? (*use first initial*): _____ ☐ No ☐ Don't know
If YES to 7a…	
7b. What kind of cancer did that person have? (*initial: type of cancer*)	
7c. What was that person's relationship to you? (*initial: relationship*)	
7d. Were they a blood relative? (*initial: Y/N/IDK*)	
7e. What was their sex? (*initial: M/F*)	
7f. How old were they when they died? (*initial: age at death*)	
7g. What year did they die? (*initial: year*)	
8a. Has anyone in your immediate family had cancer and died, who we haven't already talked about? This includes your parents, siblings, spouse, and children.	☐ Yes If YES, who? (*use first initial*): _____ ☐ No ☐ Don't know
If YES to 8a…	
8b. What kind of cancer did that person have? (*initial: type of cancer*)	
8c. What was that person's relationship to you? (*initial: relationship*)	
8d. Were they a blood relative? (*initial: Y/N/IDK*)	
8e. What was their sex? (*initial: M/F*)	
8f. How old were they when they died? (*initial: age at death*)	
8g. What year did they die? (*initial: year*)	
8h. Did they live in St. John the Baptist Parish? (*initial: Y/N/IDK*)	
If YES to 8h: **8i.** What city? (*initial: city*)	

9a. Has anyone in this household ever had a miscarriage?	☐ Yes If YES, who? (*use first initial*): _____ ☐ No ☐ Don't know
If YES to 9a…	
9b. When did the miscarriage(s) happen? (*initial: year*)	
9c. At what stage(s) of pregnancy did the miscarriage(s) happen? (*initial: week or month*)	
9d. Did that person live in St. John the Baptist Parish at the time of the miscarriage(s)? (*initial: Y/N/IDK*)	
If YES to 9d: **9e.** What city? (*initial: city*)	
10a. Has anyone in this household ever had a stillbirth (loss at 20+ weeks)?	☐ Yes If YES, who? (*use first initial*): _____ ☐ No ☐ Don't know
If YES to 10a…	
10b. When did the stillbirth(s) happen? (*initial: year*)	
10c. Did that person live in St. John the Baptist Parish at the time of the stillbirth(s)? (*initial: Y/N/IDK*)	
If YES to 10c: **10d.** What city? (*initial: city*)	
11a. Do any children in the household suffer from nosebleeds?	
If YES to 11a…	
11b. Who suffers from nosebleeds? (*Use initials*)	
11c. In the past month, how many nosebleeds did they have? (*Write number next to initials*)	
12a. Do any children in the household suffer from headaches?	
If YES to 12a…	
12b. Who suffers from headaches? (*Use initials*)	
12c. In the past month, how many headaches did they have? (*Write number next to initials*)	

Now I'm going to ask you some questions about yourself.

13. How would you rate your current overall health?	☐ Very good ☐ Good ☐ Fair ☐ Poor ☐ Very Poor
14a. In the past 12 months, have you visited a doctor or other health care provider for treatment or consultation about a medical condition?	☐ Yes ☐ No
If YES, to 14a... 14b. Approximately how many times?	

15. In the past month, how often did you experience the following symptoms?

	Never	1 day per week	2-3 days per week	4-5 days per week	6-7 days per week
a. Achiness					
b. Chest pain					
c. Cough					
d. Difficulty breathing					
e. Dizziness					
f. Eye pain or irritation					
g. Fatigue/lethargy					
h. Headaches					
i. Heart palpitations*					
j. Itchy skin					
k. Joint pain					
l. Light headedness					
m. Nosebleeds					
n. Skin rash or irritation					
o. Sneezing					
p. Sore/hoarse throat					
q. Watery eyes					
r. Weakness					
s. Wheezing					
t. Other: _____					

*Palpitations are when you feel like your heart is beating too hard, too fast, skipping a beat, or fluttering.

Stanford University

Now I have a few questions about the environment near your home.

16. How concerned are you about pollution in your community?	☐ Not at all concerned ☐ Slightly concerned ☐ Moderately concerned ☐ Extremely concerned
17. How often do you smell chemical odors while <u>inside</u> your home?	☐ Never ☐ A few times per year ☐ A few times per month ☐ A few times per week ☐ Daily
18. How often do you smell chemical odors while <u>outside</u> your home?	☐ Never ☐ A few times per year ☐ A few times per month ☐ A few times per week ☐ Daily

The next few questions I'm going to ask are about whether or not you work or have ever worked at an industrial facility. The reason we ask these questions is to get a sense of any potential exposure to chemicals as a result of your workplace.

19a. Does your job involve working on the property of an industrial facility or plant? (*It doesn't matter whether you're employed by the facility itself, by a contractor of the facility, or by a servicing company – only whether you work on the site of an industrial facility.*)	☐ Yes ☐ No ☐ Don't know
If YES to 19a...	
19b. How long have you worked on the property of an industrial facility?	☐ Less than one year ☐ _____ year(s)
19c. Approximately how many hours per week do you work on the property of an industrial facility?	
If NO to 19a...	
19d. Has your job ever involved working on the property of an industrial facility or plant? (*It doesn't matter whether you were employed by the facility itself, by a contractor of the facility, or by a servicing company – only whether you worked on the site of an industrial facility.*)	☐ Yes ☐ No ☐ Don't know

If **YES** to 19d...

19e. How long did you work on the property of an industrial facility?	☐ Less than one year ☐ ___ year(s)
19f. Approximately how many hours per week did you work on the property of an industrial facility?	

Now I'm going to ask you a few short questions about tobacco use.

20. How often does <u>anyone</u> smoke inside your home? Would you say daily, weekly, monthly, less than monthly, or never?	☐ Daily ☐ Weekly ☐ Monthly ☐ Less than monthly ☐ Never ☐ Don't know
21a. Altogether, have you smoked at least 100 or more cigarettes, cigars, or other tobacco products in your entire lifetime?	☐ Yes ☐ No ☐ Don't know

If YES to 21a...

21b. For how many years have you smoked?	☐ Less than one year ☐ ___ year(s)
21c. How many days per week did you smoke in the last month?	☐ 7 days per week ☐ 2 to 6 days per week ☐ 1 or fewer days per week

22. Finally, are there any other relevant health or environmental issues that we haven't talked about that you think we should know?

Thank you for your time!

Appendix 2

Code Abstract

```r
# `residents` refers to the dataframe containing one row per resident represented in the survey.

# the lookup() function returns the corresponding SEER prevalence stat for the given race/age/sex input.

# This arbitrary seed has been set for all Monte Carlo calculations.
set.seed(140637)

# setting loop to repeat simulation 10,000 times.
for(i in 1:10000) {
   sim <- c() # creating / resetting an empty vector to store the next simulated values.

# setting loop to run calculation for each resident (i.e. each row in `residents` dataframe).
for(j in 1:nrow(residents)) {
# retrieving relevant SEER prevalence stat as a decimal.
   x <- lookup(residents$race[j], residents$age[j], residents$sex[j])

# assigning a resident a simulated binary cancer diagnosis (1, cancer; 0, no cancer) using their SEER stat (x) as
probability.
   sim[j] <- sample(c(0,1), size = 1, replace = TRUE, prob = c(1-x, x))
}
# the vector of simulated resident cancer diagnoses are saved to be compiled (cbind()) with the others.
}

# The final result gives a data frame with one row per resident, along with a column per simulation (10,000), each cell
containing
# either 0 or 1 based on the sampled value. The sum of each column divided by the number of rows then gives the
cancer prevalence
# for the simulation. These 10,000 simulated prevalences naturally give a normal distribution with the median
simulated prevalence
# at its center. P-values are then calculated by the number of simulated prevalences >= the survey population's
cancer prevalence,
# divided by the number of simulations (10,000).
```

Appendix 3

Demographic Breakdown of Survey Population and Corresponding SEER Probabilities

Race	Sex	Age	Total Sample Size	Zone 1 Sample Size	Zone 2 Sample Size	SEER Probability
Black	M	0-9	85	49	36	0.0692%
Black	M	10-19	104	48	56	0.1382%
Black	M	20-29	82	54	28	0.2256%
Black	M	30-39	65	37	28	0.4453%
Black	M	40-49	65	40	25	1.1497%
Black	M	50-59	89	58	31	4.1103%
Black	M	60-69	84	54	30	12.8086%
Black	M	70-79	46	23	23	24.8125%
Black	M	80+	11	8	3	29.4374%
Black	F	0-9	71	40	31	0.0634%
Black	F	10-19	107	55	52	0.1352%
Black	F	20-29	56	28	28	0.2442%
Black	F	30-39	80	48	32	0.7119%
Black	F	40-49	90	45	45	2.0842%
Black	F	50-59	91	59	32	4.6132%
Black	F	60-69	79	50	29	8.3256%
Black	F	70-79	59	35	24	11.8842%
Black	F	80+	28	20	8	12.1149%
White	M	0-9	8	0	8	0.0947%
White	M	10-19	10	0	10	0.2258%
White	M	20-29	12	1	11	0.4143%
White	M	30-39	10	1	9	0.7700%
White	M	40-49	10	1	9	1.5339%
White	M	50-59	14	1	13	3.8687%
White	M	60-69	18	1	17	10.3809%
White	M	70-79	9	3	6	21.9162%
White	M	80+	7	1	6	29.0692%
White	F	0-9	7	0	7	0.0909%
White	F	10-19	10	0	10	0.2003%
White	F	20-29	11	1	10	0.4296%
White	F	30-39	14	0	14	1.1432%
White	F	40-49	5	0	5	2.9117%
White	F	50-59	9	2	7	5.9617%
White	F	60-69	16	3	13	10.3736%
White	F	70-79	15	5	10	15.3738%
White	F	80+	5	2	3	16.9960%
Hispanic	M	0-9	2	0	2	0.0842%
Hispanic	M	10-19	4	1	3	0.1992%
Hispanic	M	20-29	1	1	0	0.3187%
Hispanic	M	30-39	2	0	2	0.5131%
Hispanic	M	50-59	1	0	1	2.3494%
Hispanic	M	80+	1	0	1	21.1148%
Hispanic	F	0-9	6	1	5	0.0749%
Hispanic	F	20-29	2	1	1	0.2970%
Hispanic	F	30-39	2	0	2	0.7641%
Hispanic	F	40-49	1	0	1	1.9653%
Hispanic	F	50-59	2	0	2	4.1819%
Hispanic	F	70-79	1	0	1	9.9048%
TOTAL			1507	777	730	
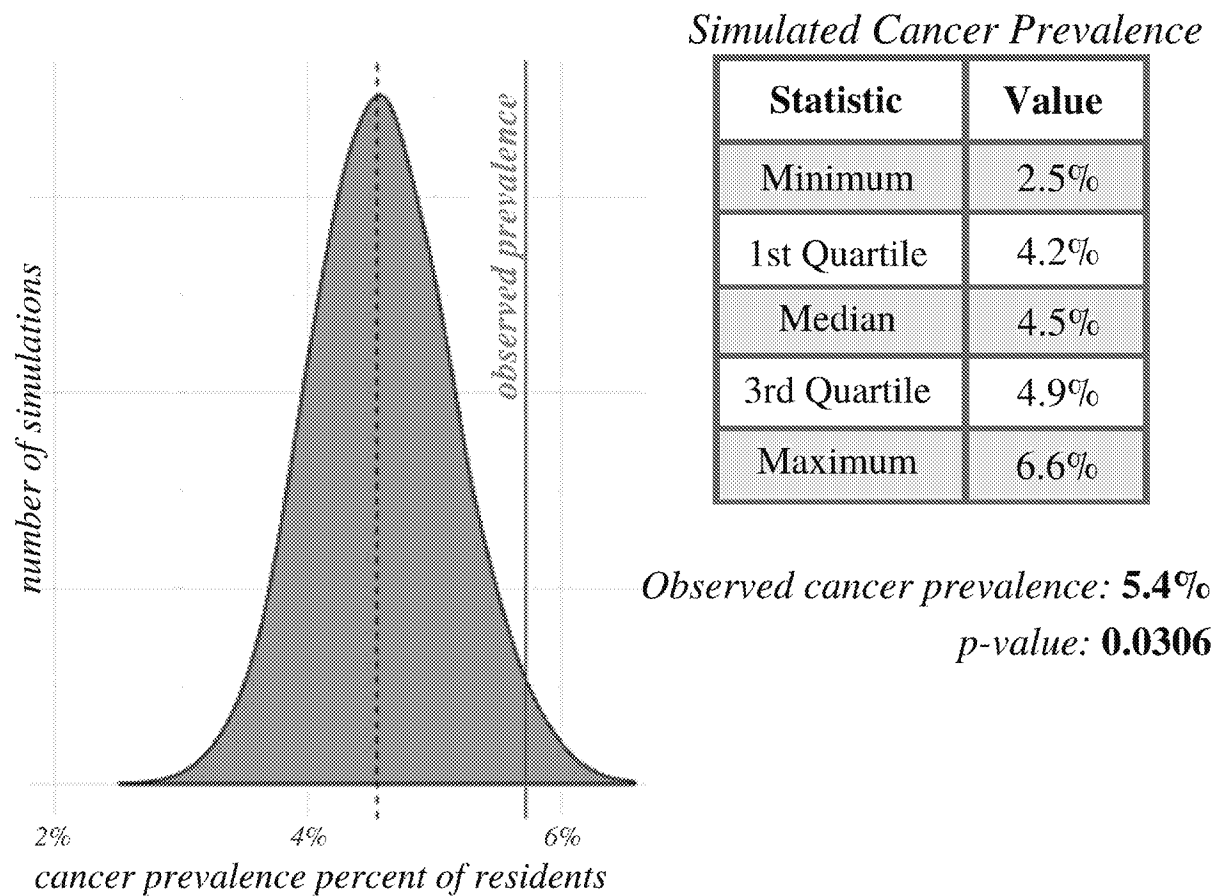
Simulated and observed 23-year cancer prevalence*

Simulated Cancer Prevalence

Statistic	Value
Minimum	2.5%
1st Quartile	4.2%
Median	4.5%
3rd Quartile	4.9%
Maximum	6.6%

number of simulations — *cancer prevalence percent of residents* — *observed prevalence*

Observed cancer prevalence: **5.4%**

p-value: **0.0306**

*after removing all residents who live in households
where anyone smokes on a daily basis

Simulated and observed 23-year cancer prevalence by zone*

number of simulations

observed prevalence Zone 1

cancer prevalence percent of residents
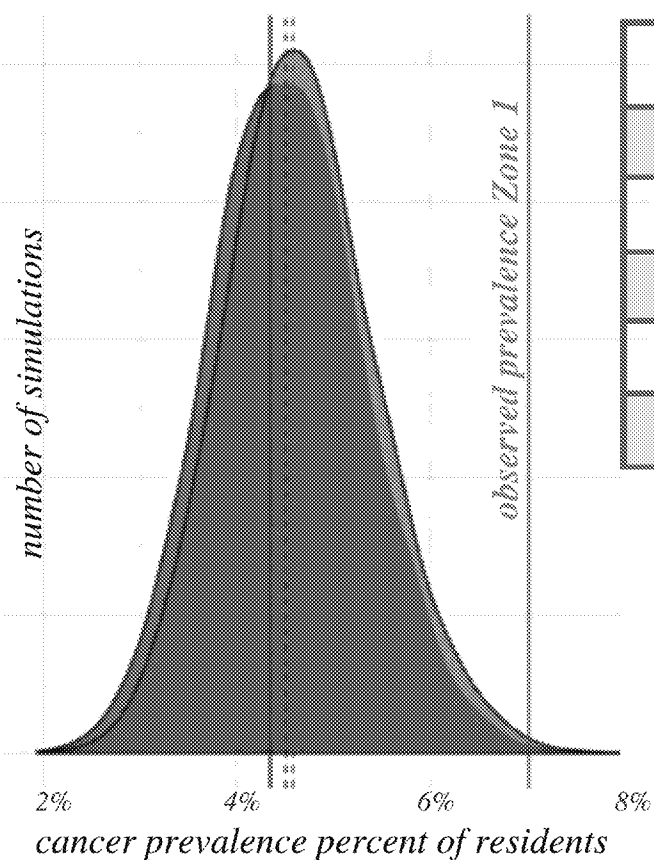
Simulated Cancer Prevalence

Statistic	Zone 1 value	Zone 2 value
Minimum	2.0%	1.9%
1st Quartile	4.1%	4.0%
Median	4.6%	4.5%
3rd Quartile	5.2%	5.0%
Maximum	8.1%	7.6%

Observed cancer prevalence:
Zone 1: **7.0%**; Zone 2: **4.3%**
p-value:
Zone 1: **0.0032**; Zone 2: **0.6139**

*after removing all residents who live in households where anyone smokes on a daily basis